

**The Benefits and Limitations of Student Evaluations of Professors  
in the Ball State University College of Business:**

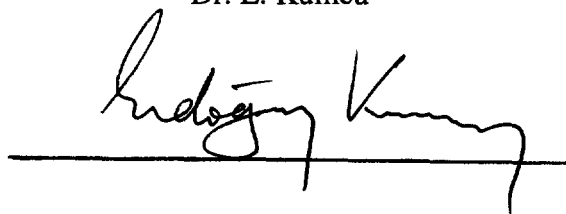
**As Discussed by Students**

**An Honors Thesis (HONRS 499)**

**By**

**Katherine D. Bowen**

**Dr. E. Kumcu**



**Ball State University**

**Muncie, Indiana**

**April, 1999**

**May 8, 1999**

Spec.  
Tech.  
20  
2000  
2000  
1999  
Ball

This thesis is a study about the opinions of students on student evaluations of professors in the Ball State University College of Business. These opinions and feelings were gathered in two focus group “discussion” sessions. The findings from these sessions were then compared to secondary research gathered by the researcher. These findings were then used to develop conclusions and recommendations concerning the student evaluation process and forms used in the Ball State University College of Business.

## **Table of Contents**

### **Summary:**

Short Statement of Objectives.....	I
Short Statement of Methods.....	II
Short Statement of Findings.....	II-IV
Short Statement of Conclusions and Recommendations.....	IV-VI

### **Introduction:**

Purpose.....	1-2
Objectives.....	2-3
Secondary Research as Related to Objectives.....	3-14

### **Research Methodology:**

Type of Study.....	14-15
Focus Groups.....	15-16
Intent of Study.....	16-17
Definition of Population and Sample.....	17
Sample Design and Technique.....	17-18
Data Collection Method.....	18-19

Determining Sample Size.....	19
------------------------------	----

## Limitations:

Sample Size Limitations.....	19
Sample Selection Limitations.....	19-20

## Analysis of Findings:

Types of Analysis Chosen.....	20-21
Findings as Related to Objectives.....	21-27
Conclusions and Reccomendations...	27-30

## Bibliography

## Appendices:

Focus Group Plan.....	Appendix I
Focus Group Explanations.....	Appendix II
Potential Biases.....	Appendix III
Student Characteristics.....	Appendix IV
Student Evaluation Form.....	Appendix V
Focus Group Participants.....	Appendix VI

## **Summary**

### **Short Statement of Objectives**

The overall goal of this thesis is to explore students' feelings about the benefits and limitations of student evaluations of professors in the College of Business at Ball State University. The basic of this research objectives are:

- to understand the attitudes and beliefs about student evaluations of professors in the College of Business
- to explore the effects of student personality characteristics on the evaluation of a professor
- to explore if student or class characteristics affect student ratings
- to examine if student ratings are affected by their intended use, i.e. promotion, tenure, or course critique
- to understand the limitations of student ratings
- to understand other miscellaneous factors that could influence student ratings of an instructor
- to understand what the evaluations are developed to measure and the opinions of students on what the evaluations actually measure
- to examine opinions on the validity of student evaluations

### **Short Statement of Methods**

In this research two methods were used. First, the researcher conducted secondary research. This research was used in the development of objectives and as a basis for development of the focus group discussion guide. Once the secondary research was completed the researcher developed the research objectives. Next, a discussion guide for the focus groups was developed.

The participants for the focus groups were selected in a non-random manner. There were several screening qualifications that the members of the target groups must meet in order to qualify to be a participant in the focus group. Once the participants were selected the focus groups were conducted. An instant analysis followed directly after the focus groups. Then, at a later time, the researcher conducted a cut and paste analysis. This analysis involves referring to a taped recording of the groups and drawing conclusions from them. This analysis allowed the researcher to find connections between the responses of the individuals in the focus groups.

### **Short Statement of Findings**

The findings of the focus groups allowed the researcher to understand the opinions of the students (focus group participants) on the benefits and limitations of student evaluations. The research showed that overall the participants felt there were problems with the evaluations. However, they were unable to overcome their objections by finding a suitable replacement for current evaluations. When related to the objectives, the results were as the

researcher expected.

Objective 1 was to understand the attitudes and beliefs about student evaluations of professors in the College of Business. The participants felt that there were many problems with the evaluations as they stand now. If some simple changes were made they could be very beneficial to both professors and students.

of teaching and class development skills.

Objective 2 was to explore the effects of student personality characteristics on their evaluation of a professor. The participants felt that generally personality characteristics had little, if any, effect on how they would rate a professor.

Objective 3 was to explore if the student or class characteristics affect student ratings. The participants once again felt that these characteristics did not affect the ratings they give a professor. They felt there were other factors that have a greater effect on the evaluation scores than those in the list they were given to choose from.

Objective 4 was to examine if student ratings are affected by their intended use. Most of the participants felt that the rating given would depend on the situation and the professor. Most participants also stated that they would rate an instructor equally no matter what the intended use of the evaluations are.

Objective 5 was to understand the limitations of student ratings. There were many limitations listed, including the length of the evaluation, the style of the evaluation, and the timing of the evaluation. The participants felt the benefits outweighed the limitations.

Objective 6 was understanding other miscellaneous factors that could influence student

ratings of instructors. The participants were given a few examples to focus on. The participants actually felt that the chosen examples had very little to do with how they would rate an instructor. They offered their own factors that they felt actually influenced the evaluation more. Many of the participants stated that more than any other factor their evaluation score reflects whether or not they like the instructor. The participants also stated that a professor's teaching style had little to do with whether they liked a teacher or not. The participants felt that how well the professor related to their students had much more to do with whether students liked them or not.

Objective 7 was to understand what the evaluations are developed to measure and the opinions of students on what they actually measure. The participants felt that they were not sure what the evaluations were developed to measure. They felt the evaluations actually measure the popularity of the instructor. They also stated the evaluations measure how well the instructor is liked rather than the teaching effectiveness of the instructor or the structure of the course.

Objective 8 was to examine opinions on the validity of student evaluations. The participants felt that there were several things that could affect the validity of the evaluations. The participants felt that overall the evaluations had some validity problems. However, the participants believed the evaluations were useful to the professors in development of teaching and class development skills.

### **Short Statement of Conclusions and Recommendations**

The researcher has come to the conclusion that there are limitations to the evaluation



form and style used in the College of Business. However, there are no other “forms” that could provide the information in such a timely and organized manner. The researcher has also come to conclude that the participants also feel the evaluations are beneficial and useful for both the professors and the students.

The researcher has many recommendations for the improvement of the student evaluations. Although it is not feasible to conduct a series of evaluations throughout the semester, it may be beneficial to the students and professors to conduct at least one other evaluation earlier in the semester. If the evaluation was offered earlier in the semester the professor may be able to make changes to the structure and/or content of the class. This process would also allow the students to feel more empowered in their education. Another recommendation may be to use students as part of the team that develops the student evaluations. The participants in this research all had very powerful and feasible ideas for the improvement of the evaluations. This would also allow the student to become more empowered in the educational process.

If the researcher had to choose a few major recommendations for the improvement of the student evaluation process there would be two. The first is to give to students more power, and the second is to give more emphasis and weight to the evaluation scores in promotion and tenure decisions. Students are in the classroom every day and they know what makes an popular instructor and what does not. Often if an instructor is considered popular then students equate this popularity with the effectiveness of the instructor. If professors know that most promotion and tenure decisions come from their research and service work they will not worry about how well they teach to and interact with their students.

The researcher feels that the evaluations should measure both instructor popularity and teaching effectiveness. There is no way for the two dimensions to be effectively separated in the classroom or an evaluation form. Therefore, the evaluations should be used to view the students' actual opinions of their professors. These opinions often include feelings about teaching effectiveness and popularity.

# **Introduction**

## **Purpose**

The purpose of this thesis was to evaluate students' opinions and beliefs about the usefulness and effectiveness of student evaluations in the College of Business at Ball State University. This research looked at the factors that influence a students' evaluation of an instructor. Finally, this research showed what improvements could be made in the current College of Business student evaluations and their use.

This is an important topic to research and understand in relation to both students and faculty. These evaluations are used in course adaptation, and promotion and tenure decisions. There have always been questions and doubts as to the validity and usefulness of these evaluations. This thesis, with its conclusions and recommendations, provides a useful source of information to those in the College of Business responsible for the decisions made in relation to student evaluations.

This thesis contains primary and secondary research. The secondary research was gathered from educational journals and internet articles. The primary research was gathered during two focus groups.

This thesis contains several sections. The first is an overview, or summary, of the entire thesis. Next, the objectives are discussed. Third, the secondary research is related to the objectives. Next, the research methodology is discussed. This section contains: the type of study, an explanation of focus groups, the intent of the study, a definition of the population and sample, the sample design and technique, the data collection method, and a determination of

sample size. The fifth section discusses the limitations of the research. The limitations include: sample size limitations and sample selection limitations. The next section focuses on the types of analysis chosen and the analysis of findings. In this section the findings are explained as related to the objectives. The final section covers the conclusions and recommendations of the researcher. An appendix is added to allow for further explanation of several parts of the thesis.

### **Objectives**

The objectives of this research were several. The overall goal was to determine the benefits and limitations of student evaluations of professors in the College of Business. These objectives are focused on the opinions of students toward these evaluations. The specific objectives of this research were as follows:

- Objective 1: to understand student attitudes and beliefs about student evaluations of professors in the College of Business,
- Objective 2: to understand the attitudes and beliefs of students concerning the questions asked in the student evaluations,
- Objective 3: to explore the effects of student personality characteristics on the evaluation of a professor,
- Objective 4: to explore if class characteristics affect student ratings,
- Objective 5: to examine if student ratings are affected by their intended use, i.e. promotion, tenure, or course critique,

- Objective 6: to understand the limitations of student ratings,
- Objective 7: to understand other miscellaneous factors that could influence student ratings of an instructor,
- Objective 8: to understand what the evaluations intend to measure and the opinions of students on what the evaluations actually measure, and
- Objective 9: to examine the validity of student evaluations.

### **Secondary Research as Related to Objectives**

The first of the research objectives, to understand the attitudes and beliefs about student evaluations of professors in the College of Business, is designed to act as an overview to the remainder of the objectives of this research. This objective focuses on the overall opinions of students about the student evaluations. In his article titled “Course Evaluations Important When Fair,” Duane M. Gran (1998, p.1) states the following:

“The purpose behind evaluating professors is to foster a responsive and evolving academic environment. Without student evaluations it is very easy for the administration to disregard student complaints as isolated incidents. Teacher evaluations are valuable, but they also have drawbacks.

The main problem with the evaluation system is that students, like most people, are not impartial. Clearest evidence of bias is when students claim they ‘earned’ a good grade, but in the same breath claim another professor ‘gave’ them a poor grade. As students with self interest, we often evaluate a teacher based on how easily we can get a good grade, or not have to work to avoid a bad grade.

Given student’s often arbitrary use of the teacher evaluation process, it is no wonder that many professors dislike the system. The professor wishes to maintain high standards, but if they do so many students will penalize them with unfair evaluations.

I believe that this is the opinion many students hold in relation to the evaluations and their use.

In another article “Student Evaluation of Teachers: Effects of Grading at College Level,” the author, Shengming Tang (1999, p.83) of Western Illinois University, stated that in his preliminary secondary research many students regarded the evaluations as “the subjective criterion of teacher effectiveness, arguing that students rate most highly instructors from whom they learn the least.” If this is the case then it leads to the following questions. Why do students rate some teachers higher than others? What criterion do they actually use to rate their instructors? For this reason, I feel that this objective is only an overview or basic opinion that will open doors to further paths of study.

The second of the research objectives is to understand the attitudes and beliefs of students concerning the questions asked in student evaluations. This focuses on the basic structure of the evaluations. Are there questions that are more useful than others? Are there questions that are not at all appropriate to the objectives of the evaluations?

In a basic student evaluation form there are two evaluative types of questions. The first looks at students’ reactions to the traits and behaviors of instructors, characteristics of course materials, and the social and physical environments of the class. The second type of question focuses on the students’ opinions on the outcomes and the educational content of the class.

Often it is the second type of question that causes concerns about the validity and

usefulness of evaluations. These evaluations must be able to evaluate the instructors responsibility in those areas in which the students are evaluating them. However, that responsibility is nearly impossible to measure.

Questions on the evaluations can be either open- ended or multiple choice rating questions. In his book, Student Evaluation of Instruction, Kenneth O. Doyle states that open-ended questions are often more effective because “only dialogue can provide complete information.” However, open-ended responses can not easily be “tabulated, summarized, absorbed, or normed, nor can the information be efficiently studied for its technical qualities.” Therefore, the more popular questions are of the multiple choice rating style. These are extremely popular because they can be “tabulated and studied.” (1975, p.19-23)

The choice between the styles and contents of questions allows room for differing opinions and beliefs about what types of questions should be used in the student evaluations. These differing opinions may lead to new and unique solutions to the dilemma of what styles of questions should be used and what the content of these questions should be.

The third research objective, to explore the effects of student personality characteristics on their evaluation of a professor is discussed again in Faculty Performance Appraisal by Gabbin, Cairns and Benke. (1990) The authors pose the question “do the attitudes, opinions, traits, interests, preferences and values students bring to a course affect student judgments of teaching effectiveness? In other words, are

student personality characteristics related to teacher rating form scores?” (1990, p.120) . These are important questions. If students allow their personality characteristics to cloud their judgment on the factors that constitute a “good teacher” then how can these evaluations be taken seriously. If indeed these personality characteristics do have an effect on ratings, which characteristics have an effect and which do not. These questions must be answered in order to create an effective evaluation tool.

In his book, Determining Faculty Effectiveness, John A. Centra (1979) also discusses the effect of student characteristics on the evaluation of an instructor. He lists the possible characteristics as “age, sex, college year (freshman, sophomore, and so on), academic ability, grade point average, expected grade in the course, reason for taking the course, and personality differences.” In a report by himself from 1976, he sites that the findings from that survey show “relationships between student...characteristics and student ratings were generally insignificant or small enough not to have any practical significance.” Centra felt that there were several student characteristics that had “weak or insignificant relationships” with rating effectiveness. These were: “sex, grade point average, college year, academic ability, and age” (1990, p.29). After all his research, Centra believes that these studies “indicate that student characteristics, although not systematically affecting the ratings given, may on occasion have a significant effect” (1990, p.29).

In “The Effects of Instructor Involvement on Student Evaluations of College Courses” the authors, John T. Pardeck and Sean FitzPatrick (1998, p.226), make several



definitive statements on the effect student characteristics have on their evaluation scores. They found no evidence that the sex of the student has any effect on the way the student rates an instructor. They report that “whether students are majors or non-majors does not appear to affect student ratings of instructors” and they also found that the level of the course also has no effect on the ratings the students give the instructor. (1998, p.226)

If these factors do affect the ratings a student gives an instructor then are there ways to prevent this affect? If not, can these evaluation still be considered valid and reliable? Is there a way to reduce the effect these characteristics may have on the evaluations? These are serious questions that must be addressed by faculty and administrators.

The fourth research objective, to explore if class characteristics affect student ratings, is addressed in several sources. First, in Determining Faculty Effectiveness, Centra (1990) discusses several class characteristics that may affect student ratings of an instructor. These include the type of course requirement, the subject matter area, the class size, and the method of instruction. Centra (1990, p.30) also found that “very small classes- those with less than ten or fifteen students- are most highly rated.” Those classes with between sixteen and thirty-five students received the second highest ratings. Those with more than 100 students received the third highest ratings. The classes with between thirty-five and 100 students received the lowest ratings. Centra (1990, p.31) also found that “students give slightly higher ratings to their majors or electives than to courses taken to fulfill a college requirement.”

In Evaluating Teaching Effectiveness , Braskamp, Brandenburg, and Ory (1984) discuss several course characteristics that may affect ratings. They also discuss the affect these characteristics usually have on the ratings. Their studies show that students give teachers of elective courses higher ratings than those of required courses. Students in higher level classes rate the teacher higher than students in lower level classes. Students in very small and very large classes generally rate their teachers higher than students in other courses. Students in courses in science give the lowest ratings while students in courses in applied life studies and education give the highest ratings (1984, p.44-45).

The findings of Braskamp and Ory (1994) in Assessing Faculty Work support their previous findings. These findings include: “ratings in elective courses are higher than ratings in required courses”, “ratings in higher-level courses tend to be higher than in lower-level courses”, “smaller classes tend to receive higher ratings...”, “...lower ratings are given to courses in the arts and humanities, biological and social sciences, business, computer science, math, engineering, and physical sciences” (1994, p.177).

Although the authors stated earlier that “smaller classes tend to receive higher ratings...,” they also found disputing evidence that stated “that differences because of class size have little practical significance.” (1994, p.180)

The fifth objective, to examine if student ratings are affected by their intended use, is addressed by Braskamp and Ory (1994) in Assessing Faculty Work. In this research the authors found that “ratings are more positive if the stated use is for promotion.” (1994, p.177)

Also, in Centra's (1990) Determining Faculty Effectiveness , the author addresses this question. One would feel that students may rate their professors either higher or lower depending on the stated use (promotion, tenure, or class critique) of the evaluation. However, Centra's (1990) research shows that "students tend to rate teachers similarly under both circumstances." (1990, p.26)

The sixth objective, to understand the limitations of student ratings, is addressed in Centra's (1990) Determining Faculty Effectiveness. The author discusses several limitations of the student evaluation process. The first limitation is that "because most student rating instruments elicit numerical responses that can be scored and quantified, it is easy to assign them a precision they do not possess." (1990, p.45) Due to the fact that these findings are often easily quantified, they are often assigned an importance that is much greater than they actually possess.

A second limitation is the fact that professors may often attempt to manipulate the findings of these evaluations. They try to "influence ratings but not student learning" (1990, p.45). Professors often attempt to win over their students and influence their evaluation ratings in many ways. Professors may invite students to their homes, take students out, or attempt to use other types of manipulating behavior. Professors may also assign less homework and grade more leniently in an attempt to influence better student ratings.

A third limitation comes from the educational institution and not the professor or student. Often educational institutions refer to their student evaluation program as proof

of their commitment to improving the education their students are receiving. However, many times these same institutions “do little else to help teachers develop their skills” (1990, p.45). This is not a limitation in all cases. Some professors are able to develop and learn from these evaluations and improve their teaching. For those teachers that are unable to develop and improve from evaluations alone, there need to be other forms of faculty and instructional development services. Often these do not exist. This then becomes a limitation.

In his article “Course Evaluations Important When Fair”, Duane Gran (1998, p.1) he makes an important point about the overall limitation of the student evaluation system. He states “at present, we have an evaluation system which students abuse and educators ignore. Students are clearly to blame for self-serving evaluations rooted in laziness, but professors are also to blame for flexing the power of tenure,” (1990, p.1).

The seventh objective, to understand other miscellaneous factors that could influence student ratings of an instructor, is first addressed in “Student Evaluation of Teachers: Effects of Grading at College Level” by Shengming Tang (1998). The first miscellaneous factor to be discussed is that of grades. Tang believed older research that had found a relationship between grades and student evaluations was outdated. In his new research he showed this to be true. There is a relationship between grades and evaluation scores. However, “it was not the grade, but the difference between the expected and received grades, that biased student evaluations of their professors” (1998, p. 83). Tang (1998) found that the “expected success had no relationship to evaluations

but failure to meet expectations did.” (1998, p.83)

Also, in Assessing Faculty Work by Braskamp and Ory, (1994) the authors found that “students expecting high grades rate their instructors higher than do students expecting low grades.” (1994, p.178)

Another miscellaneous factor that may affect the ratings of professors is the way in which the professor presents the class. Does a professor who entertains rather than teaches receive higher ratings than the professor who actually teaches? In the article “Measuring Teaching Performance” Edward M. Gramlich and Glen A. Greenlee (1993) stated that in their research they found “effective drillmasters are not very popular teachers...effective teachers are so because they encourage learning or awareness...” (1993, p.12).

Also, in “Monitoring and Improving Instructional Practices (and are Student Evaluations Valid?),” Bryan W. Griffin and Harbinson Pool (1998) found that in classrooms where “for every category of instruction rated, those students exposed to the more enthusiastic lecture rated the instruction and course higher” (1998, p.2). In another study with two classes of students where “both groups of students were exposed to identical tests, syllabi, texts, course goals, and grading practices, yet the group that received the more enthusiastic lecture rated each of these higher” (1998, p.2). Overall, the authors of this article found that the more enthusiastic the instructor, the higher the rating he/she receives in the student evaluation (1998, p.2).

Another miscellaneous factor that may influence the evaluation process is the way

in which the evaluation is administered. Many studies have shown that the presence of the instructor in the room during the evaluation process may have an affect on the students ratings of the instructor. In the article “The Effects of Instructor Involvement on Student Evaluations of College Courses,” John T. Pardeck and Sean FitzPatrick (1998) ask the question “does the presence of the instructor in the classroom during the course evaluation affect student ratings?” (1998, p.227). Their answer is whether or not the instructor is present in the classroom has “no statistically significant effect on how students rate a course” (1998, p.228). There is no statistically significant impact “associated with the instructor versus a designated person collecting the student evaluations” (1998, p.228). However, in Assessing Faculty Work, Braskamp and Ory, (1994) found that ratings tend to be higher if the instructor remains in the room during the administration of the evaluation (1994, p.177).

The eighth objective, to understand what the evaluations are developed to measure and the opinions of students on what they actually measure is addressed in Evaluating Teaching Effectiveness by Braskamp, Brandenburg, and Ory (1984). The authors feel that “students are appropriate sources when they are describing or judging student- instructor relationships, their views of the instructor’s professional and ethical behavior, their workload, what they have learned in the course, fairness of grading, and instructor’s ability to communicate clearly” (1984, p.36-38).

Also, in Determining Faculty Effectiveness, Centra (1990) lists these as the ten most frequently studied characteristics: communication skills, favorable attitudes to

students, knowledge of subject, good organization of subject matter and the course, enthusiasm, fairness, flexibility, encourages students to think for themselves, and good speaking ability (1990, p.18-19).

However, in Student Evaluation of Instruction, Doyle (1975) states that the measurement of these items may be affected in several ways, including the “prior experience of the rater” (1975, p.34). This prior experience consists of experience with previous instructors and people in general. The raters’ expectations, and rumors or reputations concerning the instructor may also affect the rating. These may be the characteristics that these evaluations actually measure.

The final research objective, to examine opinions on the validity of student evaluations, is discussed by Gaggin, Cairns, and Benke (1990) in their book Faculty Performance Appraisal. The two types of validity used to evaluate student evaluations are convergent and divergent validity. Convergent validity, the type typically used to evaluate student evaluations, focuses on the relationship between student and instructor ratings on the same factor or question. The results of their study on the validity of student evaluations supports the convergent validity of student evaluations.

The findings when using divergent validity are somewhat different. “Divergent or discriminant validity was assessed by seeking the answers to two related questions....is the student-instructor agreement on an evaluation factor specific to that factor, or can it be explained in terms of generalized agreement common to all the different factors?” (1990, p.115) The divergent validity also looks for the presence of a “halo effect” or do

the answers “represent true relationships among underlying dimensions?” (1990, p.115). Their study found very clear answers to these questions of divergent validity. “Student-instructor agreement on each evaluation factor is specific and distinctive from other factors. Whereas correlations between student and instructor ratings on the same factor are uniformly high, correlations between their ratings on different factors are generally low. Finally, the results in this research, although not as clear as some other conclusions, show that “there is at least some halo effect in the student ratings.” (1990, p.115). However, in their evaluations of instructors’ self-evaluations the authors found little evidence of the same halo effect.

## **Research Methodology**

All of the information discussed in the Research Methodology section is addressed in Contemporary Marketing Research by Carl McDaniel Jr. and Roger Gates (1993).

### **Type of Study**

This study of the opinions and beliefs of students in the Ball State University College of Business about the student evaluations of faculty is qualitative in nature. The data that was collected will not be quantified or used in any quantitative analysis. This



data is purely opinions and beliefs. The collected data allowed the researcher to understand the in-depth feelings and motivations of the students in the Ball State College of Business.

This research was an exploratory study into the feelings and opinions of the students about the usefulness and effectiveness of the student evaluations of professors in the College of Business. By using two focus groups of eight student participants, the researcher was able to explore and understand their motivations in their ratings of instructors. This study utilized in- depth interviews with some of the professors that were involved in the development of the evaluation form used in the College of Business. These interviews enabled the researcher to formulate objectives and create a discussion guide for the focus groups.

### **Focus Groups**

The primary research method used in this research was focus groups. The researcher followed a very strict procedure for the focus groups. First, the type of focus group to be used for this study was an experiencing focus group. This type of group discussion allowed the researcher to observe and listen to how the participants thought and felt about the evaluations. The researcher was able to use the focus groups to bring together groups of students to discuss the evaluations. These groups allowed the participants to discuss the evaluations rather than just answer questions about the evaluations.

Next, the setting for the focus groups was chosen. Due to time and scheduling constraints these focus groups were conducted in the home of the researcher. The groups were audio taped for the future reference of the researcher.

Third, participants were recruited. This tied into the sampling method and sampling frame, which will be explained later. The researcher decided that each focus group must contain eight participants. These participants must meet all the screening criteria. The researcher decided to put a time limit on the group discussions. The general rule of thumb is that no focus group should last more than two hours. The researcher placed a one and a half hour limit on the discussions.

Next a moderator was selected. The moderator leads the focus group by asking predetermined questions and guiding the conversation. The moderator for these focus groups was the researcher. Finally, a discussion guide was developed for the moderator. A discussion guide is an outline of the topics to be covered during the session. This guide was based on the research objectives and served as a checklist to make sure all the topics were covered in a certain sequence.

### **Intent of Study**

The intent of this study was to understand the benefits and limitations of student evaluations of professors in the Ball State University College of Business. It also studied the factors that may affect those evaluation scores. The other intent of this study was to probe students and explore their opinions and beliefs concerning the form of the

evaluations, the factors that influence the evaluation scores, and the use of these evaluations in promotion and tenure.

### **Definition of Population and Sample**

The population of a study is defined as the total group of people from whom information is needed. Defining the population involves defining the target groups for the study. In order to define the population or target group, the researcher must first understand whose opinions are needed to fulfill the objectives of the research. There was one basic target group in this study. The target group was students in the College of Business. These students were male or female and of any grade level. The qualifications for the students are that they must have a College of Business major or minor and have taken one or more student evaluations before the time of the focus group.

Individuals who were excluded from this study are those that have never been involved in the student evaluation process in the College of Business in any way. Any other individual may be chosen to take part in the research.

### **Sample Design and Technique**

The sampling frame for the focus group participants was that they must be students with a College of Business major or minor that had taken the student evaluation at least once before the time of the focus group meeting. The sampling method was a non-probability sample using a combination of convenience and judgment samples. This

means that the researcher chose the students that she felt were best fit for the focus group from any student within the sampling frame. There was no attempt to randomize this sample. The total sample size for the focus groups was sixteen individuals. There were eight individuals in each focus group.

### **Data Collection Method**

The selection of a sampling method depends on several considerations. These often include the objectives of the study, the financial resources available, time limitations, and the nature of the subject under investigation. The major limitations to the selection of a sampling method in this study were the objectives and time limitations. For this study the researcher chose a non-probability sample. This sample includes the selection of specific individuals from the population in a nonrandom manner. The individuals were selected on the basis of convenience. The reason for the use of a non-probability sample were its advantages. Non-probability samples are less costly than probability samples. They can also be conducted more quickly than probability samples. Although these are definite advantages there are also disadvantages to using a non-probability sampling method. The disadvantages are that the researcher does not know the degree to which the sample is representative of the population from which it was drawn. Also, the results from a non-probability sample can not be projected onto the total population.

### **Determining Sample Size**

In order to select a sample size using a non-probability sample several methods can be used. These methods are budget available, rules of thumb, and number of subgroups to be analyzed. The method the researcher chose to use in this study was based on rules of thumb. The general and effective size in a focus group is from eight to twelve individuals. Due to time constraints the researcher has chosen to conduct two focus groups comprised of eight individuals. This brought the total sample size for the focus groups to sixteen individuals.

### **Limitations**

#### **Sample Size Limitations**

The limitations to sample size are relatively simple. First, due to the type of primary research chosen there is a limit to the number of individuals needed to conduct the research. Focus Groups are most effective when using between eight and twelve individuals. Any more than twelve causes confusion and the individuals are unable to communicate all their feelings and opinions in the allotted time frame. The second limitation is scheduling. Due to the busy time period in which the focus groups were scheduled it was difficult to find more than eight individuals to participate in the groups.

#### **Sample Selection Limitations**

The sample selection limitations are also relatively simple. The first limitation

was based on the screening qualifications that were stated earlier. These qualifications somewhat limited the individuals that would be able to participate in the focus groups. Therefore, only those individuals that fit the qualifications were able to be selected to take part in the focus groups. The second limitation is based on the type of study being conducted and the opinions of the researcher. Once again, because this primary research was being conducted in a focus group format there were only a certain number of participants that are needed. Due to this numerical limitation there was also a selection limitation. This selection limitation involved the opinions of the researcher. The researcher chose the participants she felt to be of the most benefit to the group out of all those individuals that may have been selected.

## **Analysis of Findings**

### **Types of Analysis Chosen**

For the analysis of the focus groups two types of analysis were chosen. The researcher chose to use an instant analysis and a cut and paste technique (McDaniel, 1993, p.204). The differences between these two techniques are the time at which they are conducted and the depth to which they analyze the material.

The first technique, instant analysis, was conducted directly after the focus group was conducted. The moderator (researcher) made notes on the focus groups. These notes were drawn from the memory of the moderator. This technique did not use a

recorded version of the group discussion to jog the researcher's memory about what had occurred. This allowed the researcher to keep track of the opinions and ideas that were in her mind right after the focus groups. This analysis allowed the researcher to make quick judgments about what had been stated during the focus groups. The researcher made notes as soon as the sessions ended. These notes included the researchers overall opinion about what had occurred in the discussions.

The second technique, cut and paste (McDaniel, 1993, p.204), was completed at a later date. This allowed the researcher to have time to reflect on what occurred and what was said during the focus group. The researcher used recorded tapes of the focus groups to refresh her memory and used these tapes to help him/her look for common threads between the responses of the participants. This also allowed the researcher to compare responses between groups. Finally, this enabled the researcher to take direct quotes from the tapes to support her statements and findings with greater strength. The researcher used this technique to complete the findings section of the thesis. Using the notes made from the instant analysis the researcher then used the tape recorded sessions to develop the findings and conclusions.

### **Findings as Related to Objectives** (refer to Appendices I and II)

Objective 1 is to understand students' attitudes and beliefs about the students evaluations of professors in the College of Business. The focus group participants' overall attitude about the evaluations seemed to be negative. Throughout most of the

discussion the participants focused on the disadvantages and limitations of the evaluations. One participant called the evaluations “useless, because neither the students or the teachers take them seriously.” When the participants were asked to discuss their likes and dislikes of the evaluations the comments were predominantly dislikes. Several dislikes were listed. These dislikes included: the length of the evaluation, when the evaluations were administered, the vagueness of the questions, and the fact that the questions were mostly multiple choice. However, there were also “likes”, though they were few. The more commonly stated “likes” included: the uses of the evaluation, the opportunity for written comments, and the anonymity. When the participants were asked to discuss problems, one stated “the scores on the evaluations have no effect on the instructors with tenure and those are the ones that usually receive lower scores.”

Objective 2 is to explore the effects of student characteristics on their evaluation of a professor. During the section of the discussion that focused on student characteristics the participants felt that overall these characteristics have little effect on the evaluation scores. When asked if they rated professors more like themselves higher or lower than other professors the participants responded unanimously. All of the participants said no, they did not. One participant said “that’s almost dumb... a lot of my favorite professors are nothing like me.” Another said “ if every professor I had was like me I think I’d go nuts.” However, there was one characteristic that the participants did agree had some affect on their rating of a professor. That characteristic was the grade they expected in the class. One participant stated “if I’m expecting a high grade in the



class that means one of two things, either the professor is good and I'm learning what I'm supposed to, or the professor is easy and I've coasted through the course without doing much, both of those are good.. of course I'm going to rate them higher."

Objective 3 is to explore if the class characteristics affect student ratings. On most of these topics the participants had differing opinions. One major topic in this section of the discussion focused on the ratings of a professor in required vs. elective courses. The participants in the groups seemed to fall into two groups on this topic. One group stated this would probably affect their ratings of the instructor. The other group felt that this would have no affect on their rating. One participant said "if I don't like the class... I'm probably not going to like the professor either." Another participant stated "it's not their fault you had to take the class and you hate it...they still are probably a very good teacher."

Another major focus in this section was class size. The participants felt that generally class size had no affect on the ratings given to an instructor. When the discussion moved onto a list of "potential biases" (Appendix III) every participant felt a different bias would have an affect on the rating of the professor. Often these opinions were extremely different. Some students felt that grading leniency would have an affect on scores, others did not. However, almost all of the participants agreed that "course workload" and "instructor popularity" have some form of affect on the ratings given an instructor.

Objective 4 is to examine if student ratings are affected by their intended use, i.e.

promotion, tenure, or course critique. During the sections of the focus groups that related to this objective the participants stated that it depended on the situation and the professor. However, most agreed that they would rate a professor equally in all situations. One participant stated “every teacher has something about them you don't like,...you're just more willing to tell them that if you don't think they are going to get docked for it.”

Objective 5 is to understand the limitations of student ratings. The participants were able to list and discuss many of the limitations of the evaluations. Aside from the general dislikes the participants had for the evaluations they found other problems with the evaluations. One participant pointed out one distinct disadvantage to this form of evaluation. This participant felt that the structure could be a great limitation. The participant felt this could be a disadvantage because “students are not able to voice the reasons for their opinions on certain points and therefore the professor has to guess at why the students rated them lower or higher in certain areas.” Another limitation stemmed from question #16 on the evaluation itself (Appendix V). Most of the participants agreed that this question was not an adequate summary for questions #6-#15. The participants felt that another, more effective, “summary” question should be developed for the evaluation. Yet another limitation seemed to be the criteria on which the professors are evaluated. When the participants were asked to give other criteria they would like to see on the evaluation the list was long. Some of the criteria they listed included: the listening skills of the professor, grading and attendance policies, other rules of the professor, and the amount of structure the professor gives the class.

Objective 6 is to understand other miscellaneous factors that could influence student ratings of an instructor. The first miscellaneous factor discussed was whether or not the participants would rate a professor who entertained higher than a professor that just taught the class. Most of the participants felt it depended on the professor and the class. One participant said “it just depends on the professor. Some can talk and talk and talk and it drives you crazy. Others can do the same thing and you think its funny and it makes the time go faster... then you rate them according to that.”

Next the conversation focused on whether the participants rated professors based on teacher characteristics and what they learned in the class or the organization and structure of the class. The participants stated that they use a mixture of both. One participant said “its not really one or the other... it takes teacher characteristics, what you learned, and organization and structure in the class to make a good professor.” Most of the participants agreed with this point of view.

The final miscellaneous factor that the focus groups covered was whether the method used to collect the information would influence the rating of a professor. The participants felt that if the professor was not in the room during the evaluation this made it easier for the students to evaluate the professor honestly. One participant stated “if the professor was in the room I would probably be afraid to say what I really wanted to because he might look at it at it after I turned it in.” In general, most of the participants felt that the manner in which the evaluations are conducted in the College of Business is fair. They felt that any other method of collection would hinder the collection of valid and reliable data.

The eighth objective is to understand what the evaluations are developed to measure and the opinions of students on what they actually measure. The participants felt that they were unqualified to discuss what the evaluations were developed to measure. However, they felt they could discuss what the evaluations actually measure. Overall, the participants felt that the evaluations measure how well they like the professor. One participant stated "...you rate them higher or lower if you like them or not." That was the basic idea the participants gave during the discussion. They, as students, rate teachers they "like" higher than those they do not.

The final objective is to examine opinions on the validity of student evaluations. This objective was never directly approached during the discussions. However, the researcher was able to get a general idea about the participants opinions from their answers to other questions. During the discussions the moderator led the participants through several lists of characteristics that may have an effect on the ratings. Most of the students responded that many of these characteristics had little, if any, affect on how they would rate an instructor. However, these same participants did agree that whether or not they liked an instructor would affect their rating. They also indicated that the grades that were expected in the class may also affect the ratings. These discrepancies seem to show that overall these evaluations may not be as valid as the university would hope. These findings do agree with the findings from the secondary research. Although there are a few differences, the basic findings are the same. Overall, student evaluations are beneficial when administered properly. There will always be questions as to their validity. There is no definitive proof they are valid or invalid. As long as they serve the

purpose they are set forth to they are effective instruments of evaluation. There are limitations and problems with the evaluations in the College of Business. However, with minor adjustments and improvements the evaluations could be extremely useful to the College of Business.

Table 1

Findings as Related to Objectives

---

Objective 1:	Overall attitudes negative, focus on disadvantages and limitations
Objective 2:	Most of the discussed characteristics have little affect on the validity of evaluations, grades may affect validity of ratings, overall evaluations not as valid as they should be
Objective 3:	Student characteristics do not great affect on ratings, ratings of professor are not biased by these characteristics
Objective 4:	Participants were not unanimous on the affect class characteristics have on ratings, class size has very little affect on ratings
Objective 5:	Intended use of evaluations affects ratings only in certain situations
Objective 6:	Disadvantages in the form of the evaluation, Question #16 not an adequate summary of evaluation
Objective 7:	Several miscellaneous factors may have an affect on ratings only in specific situations
Objective 8:	Evaluations actually measure how well the student likes the professor

---

**Conclusions and Recommendations**

The researcher has reached several conclusions after the collection and analysis

of this data. First, there are several limitations to the evaluations now used in the Ball State University College of Business. However, the benefits of the evaluations outweigh the limitations. Strong benefits include giving feedback to professors, students and administrators.

The researcher has also concluded that although the evaluations fail in many areas they are necessary. These evaluations allow the students an outlet to express their opinions and ideas to the instructors and faculty. These evaluations also allow the faculty and instructors to view the opinions and feelings of students and act on them.

The final conclusion that the researcher reached was that there are many ways in which these evaluations can be improved. These improvements would benefit the students, professors and faculty. The basic format for the evaluations is beneficial. However, there are improvements that can be made.

The researcher has taken the improvements that were discussed in the focus group discussions and channeled them into recommendations. There are several recommendations. As one participant stated in the focus group discussions, it may be beneficial for instructors and students if the evaluations were administered more than once a semester. The researcher recommends that the evaluations be given in three sets. One given after three or four weeks. Another given around midterms. Then, the final evaluations should be given at finals. This method allows the instructors to view the students' comments earlier in the semester and make improvements in the class. The second evaluation would act as reference for

the instructor. This evaluation would allow the instructor to see if the improvements they made have helped the students. The final evaluation would act as an overall evaluation of the professor for the semester. This recommendation would allow the students to feel more empowered in the educational process. This would also allow the instructor to adapt the course to better benefit the students. If evaluations are given only at the end of the semester then the instructor has no basis for correction during the semester itself.

Another recommendation is to allow for more open-ended responses in the evaluations. Open-ended responses allow for fuller expression on the part of the students and fuller understanding on the part of the instructor. This would allow the instructor to look at the ratings they received and see more than numbers. This would show the instructors why students rated them the way they did in particular categories. This recommendation would give students more power in the educational process. It will also give the instructors a definite reasoning behind students ratings. The ratings would no longer appear to be quite as arbitrary as they do now.

A final recommendation would be to allow students to be part of the committee that develops the evaluations. Students have very concrete ideas of what constitutes an effective instructor. The researcher believes that the findings show that sometimes the instructors and faculty lose sight of this. There are certain attributes that students look for and respect. Allow the students to help the committee focus their ideas more clearly. This will help to develop an evaluation that students and instructors can relate to and understand.

The overall recommendation of the researcher is to give the students more power in the educational process. It is their education. The participants of the focus groups had very strong ideas and opinions about these evaluations. There is room for improvement. This improvement can be made, and it may be more effective, if the students are allowed more voice in their education.

Table 2

Major Conclusions and Recommendations

- 
- Evaluations are failing in many areas, but they are necessary
  - Benefits of the evaluation outweigh the limitations
  - There are many ways in which the evaluations can be improved
  - Evaluations should be administered more than once a semester
  - Allow for more open-ended responses in the evaluations
  - Allow students to be part of the committee that develops the evaluations
-



## **Bibliography**

- Braskamp, Larry A. Assessing Faculty Work. San Francisco: Jossey-Bass Publishers, 1994
- Braskamp, Larry A. Evaluating Teaching Effectiveness. Beverly Hills: Sage Publications, 1984.
- Centra, John A. Determining Faculty Effectiveness. San Francisco: Jossey-Bass Publishers, 1979.
- Doyle, Kenneth O. Student Evaluation of Instruction. Lexington, Mass: Lexington Books, 1975.
- Gabbin, Alexander L. Faculty Performance Appraisal. James Madison University. Center for Research in Accounting Education, 1990.
- Gramlich, Edward M. "Measuring Teaching Performance." *Journal of Research in Economic Education*, 1993.
- Gran, Duane M. "Course Evaluations Important When Fair." 1998.  
[http://ragnar.spinweb.net/dn/2\\_3\\_98.html](http://ragnar.spinweb.net/dn/2_3_98.html)
- Griffin, Bryan W. "Monitoring and Improving Instructional Practices (and are Student Evaluations Valid?)." Georgia: Georgia Southern University, 1998)
- McDaniel, Carl Jr. Cotemporary Marketing Research. Minneapolis: West Publishing Compnay. 1993
- Pardeck, John T. "The Effects of Instructor Involvement on Student Evaluations of College Courses." St. Louis: Southwest Missouri State Unviversity, 1998.
- Tang, Shengming. "Student Evaluation of Teachers: Effects of Grading at College Level." Illinois: Western Illinois University, 1999.

## **Appendix I: Focus Group Plan**

This focus group plan was used during the focus groups as a guide for the moderator/ researcher.

### **I. Warm-up**

- A. Explain topic of thesis and purpose of focus group
- B. Explain focus groups
- C. No correct answers- only your opinion. You are speaking for other students like yourself.
- D. Need to hear from everyone
- E. Please- only one person talking at a time. No side discussion- I'm afraid I'll miss some important comments.
- F. Don't feel bad if you don't understand some of the things I'll be asking you about- that's OK and important for me to know. If your views are different from others in the group that's important for me to know. Don't be afraid to be different. I'm not looking for everyone to agree on something unless they really do.
- G. This is being videotaped because I want to concentrate on what you are saying- so I don't have to take notes.

### **II. What is the first thing that comes to mind when I mention student evaluations of professors in the College of Business?**

- A. Likes? Dislikes?
- B. Problems?
- C. Changes you would like to see?
- D. How do you feel about the current College of Business student evaluations of professors?
- E. Is there something you would consider a beneficial alternative to the traditional student evaluations? What and why?

### **III. Now I would like to show you the evaluation the College of Business uses. I am interested in how you feel about this evaluation.**

#### **SHOW COLLEGE OF BUSINESS STUDENT EVALUATION**

- A. What do you like/ dislike about it?
- B. Do you see any distinct advantage or disadvantage over other forms of evaluation?
- C. Do you think Question #16 is an adequate summary of Questions #6-#15? Why or why not?
- D. Do you think that Questions #16 is a useful question in the overall evaluation of teaching effectiveness? Why or why not?
- E. If you answered "yes" to "C" or "D" why ask Questions #6-#15?
- F. What other criteria would you use to evaluate a professor? List 2 or 3.

### **IV. Now I would like to ask you some questions on specific factors that may effect student evaluations of professors.**

- A. Would you rate a professor higher or lower if you knew the use of the evaluation as

- for tenure or promotion rather than for a critique of teaching style of the class?
- B. Do you think the timing of the evaluation (during final exams vs. earlier in the semester) has an effect on the ratings?
  - D. Do you usually rate a professor higher or lower in smaller or larger classes?
  - E. Of these potential biases pick one than you feel would have the most effect on ratings and explain why: course difficulty, grading leniency, instructor popularity, student interest in subject, course workload, class size/ enrollment, required vs. elective course

V. Now I would like to ask you a few questions about student characteristics that may effect evaluations.

- A. Do you rate a professor you feel to be more like yourself higher than other professors?
- B. Do you rate a professor higher or lower if you are expecting a higher grade in the class?
- C. Do you feel that a students reason for taking a class has an effect on their rating of the professor?
- D. Of these student characteristics: age, sex, college year, academic ability and GPA; which, if any, do you think effect evaluation scores?

VI. Now I would like to ask you a few miscellaneous questions about student evaluations.

- A. Are teachers who entertain rather than teach rated higher by students?
- B. Do students tend to focus their evaluations on teacher characteristics and what they learned in the class or the organization and structure of the class?
- C. Does the method used to collect the information influence the students' evaluations of the overall teaching competence of an instructor or the quality of the course? Do you see any other alternatives to this method of collection?

VII. Now, finally, I would like to ask for any further comments or questions.

THANK YOU FOR THE COOPERATION

## **Appendix II: Focus Group Explanations**

In order to ensure that all of the information that was discussed in the focus groups is reported in this thesis the researcher has chosen to complete the analysis and explanations following the discussion guide for the focus group (Appendix I).

When the moderator asked the participants “what is the first thing that comes to mind when I mention student evaluations of professors in the College of Business?” the responses were unanimous. All the participants in both focus groups were not thrilled with the mention of the topic. They all seemed to express disgust with the process and the evaluation itself. One participant responded that he felt the evaluations were “useless, because neither the students or the teachers take them seriously.”

Then, the moderator continued on with the first set of questions. The moderator asked for the likes or dislikes of the participants pertaining to the evaluations. The participants were eager to respond. Many different opinions were expressed, mostly dislikes. Some of the commonly stated “dislikes” were:

- the length of the evaluations,
- when they are administered,
- the fact that many of the questions are very vague,
- and you are only given multiple choice responses.

Some of the more commonly stated “likes” were:

- the uses of the evaluations,
- the opportunity for written comments,

- and the fact that they are anonymous.

Next, the moderator wanted to know if the participants felt there were any problems with the evaluations. This did not draw as large a response. The participants did not know enough about the evaluation process to be able to discern problems within it. One participant, however, did feel that one major problem with the evaluations was the fact that “the scores on the evaluations have no effect on the instructors with tenure and those are the ones that usually receive lower scores.”

When the moderator asked the participants what changes they would like to see in the evaluations there were many different answers.

- One participant felt that a more comprehensive evaluation with more specific questions would be valuable.
- Another participant said that “they should be used earlier in the semester so the teacher can make changes to help the students.”
- Yet another participant felt that there should be more open ended, or written response questions, because “the multiple choice questions don’t allow students to respond freely.”

When the moderator asked if there were things that the participants considered “beneficial alternatives” to the traditional student evaluations used in the Ball State University College of Business, the participants had little to say in response. The moderator probed further. The participants stated that they did not know of other alternatives available. However, one participant came up with the idea of a continuous

evaluation process. Rather than a one time evaluation at the end of the semester. This would consist of a series of evaluations given throughout the semester instead of only one evaluation given at the end.

In the next part of the focus group discussion, the moderator presented the participants with a copy of the current evaluation used in the College of Business. The moderator then asked questions pertaining to the evaluation itself. The first question focused on what the participants liked and disliked about the College of Business evaluation. In answer to this question the participants reiterated the statements they made earlier in the group. These were comments concerning the length of the evaluation and the way in which it is administered. Their likes were somewhat different, however. They liked the fact that they were developed within the college and many of the teachers that were evaluated actually took part in the development of the evaluations.

The moderator asked if the participants saw any distinct advantages or disadvantages to this evaluation form over other forms of evaluation. The participants once again responded that they were not acquainted with many of the other forms of evaluation. However, when they were further prompted by the moderator, one participant stated that the evaluation form that the College of Business uses may have the advantage of being more structured than an evaluation that was entirely open ended questions. This would be an advantage because students are able to respond in a way that can be made into statistics. This would be easier for professors to look at and identify their strengths and weaknesses. Another participant then pointed out that the structure

could also be looked at as a disadvantage. Students are not able to voice the reasons for their opinions on certain points and therefore the professor has to guess at why the students rated them lower or higher in certain areas.

The instructor asked the participants if they felt question #16 was an adequate summary of questions #6-#15. Most of the participants responded that they looked at question #16 as a separate question and not a summary of the previous questions. Most of the participants felt that it was not an adequate summary.

- One participant said that “overall I may rate a professor a lot higher than I would rate them on the other questions because of things that have nothing to do with the other questions.”
- Another participant stated that they did not feel that question #16 was an adequate summary because “there are a lot of other things that I look for in a good professor that are not included in those questions.”

The moderator then asked the participants if they thought question #16 was a useful question in the overall evaluation of teaching effectiveness. The participants, for the most part, felt that this question was more a question of whether or not they liked the professor.

- One participant said that “when I answer that question I think about whether I like the ‘prof’ or not.”
- Another participant then said “yeah, the other questions are more about how well they teach, this one is more about what the students think of them as a person, at least

that's how I answer it.”

When the moderator then asked the participants to list two or three other criteria they thought professors should be rated on there was an explosion of conversation. All of the participants had an opinion on this question. The moderator instructed the participants to work their way around the table and have everyone list two or three other criteria. There were many different suggestions.

- One participant said that a criterion should be “if the professor is interesting while they teach the class.”
- Another said that the instructor should be rated on whether they listen to students and take their suggestions.
- Another said that a useful criterion would be whether the grading and attendance policies of the professor follow university regulations, or if the professors “just make up their own rules as they go along.” Many of the criteria the other participants listed were similar to these;
  - how interesting the professor makes the class,
  - the rules of the professor,
  - whether or not the professor shows favoritism,
  - whether the professor admits they are wrong when they are,
  - whether they give enough structure or they just expect the students to figure it out for themselves,
  - and the workload of the course.



In the next section of the focus group discussion the moderator asked the participants questions about specific factors that may affect student evaluations of professors. The moderator asked the participants if they would rate a professor higher or lower if they knew the use of the evaluation was for tenure or promotion decisions rather than for a critique of the teaching style. The participants said it really depended on the situation and the teacher. Most agreed that generally they would rate a teacher equally in both situations. One participant stated that “if it was a teacher that I hated and it was for tenure or promotion I would rip them up, but if it was a teacher I liked and it was going to be used for the same thing I would probably much nicer.” The participants seemed to feel they would rate a teacher more honestly, but not necessarily better or worse, if the stated use was for a critique of the teaching style. One participant said “every teacher has something about them you don’t like,...you’re just more willing to tell them that if you don’t think they are going to get docked for it.”

The moderator asked the participants if they thought the timing of the evaluation had any affect on the ratings they give the teacher. There was a range of responses. Some participants felt they would rate a professor the same no matter when they took the evaluation. Others felt that at certain times in the semester they would rate a professor higher or lower depending on the circumstances. One participant said “if I have to fill out an evaluation right after I get back a midterm that I got a C on when I thought I should have gotten a B, I may not be as nice as I would have at another time.” On the other hand, another participant said that “you know if a professor is good or not, just

because you got a C instead of a B should have nothing to do with how you rate them.”

Next, the moderator asked if the participants rate teachers higher or lower depending on the size of the class. The participants felt that usually the size of the class has nothing to do with the ratings of the instructor. One participant stated that “one of the best professors I’ve ever had was in a 300 student Western Civ lecture, the other was in a 15 student Humanities class.” Another participant said “...that doesn’t matter, a teacher is either good or bad, it doesn’t matter how many people are in the class.”

The moderator asked the participants to pick from a list of “potential biases.” (Appendix III) Almost every participant had a different answer concerning the one they felt would have the most affect on the ratings an instructor would receive. Some felt that “grading leniency” was the strongest of the biases listed, others felt that teacher’s grading policies had nothing to do with the evaluation scores. Most of the participants did agree that “course workload” and “instructor popularity” often had a great affect on the way they rated the professor. However, they also said that they felt there were other things would have a greater affect on the ratings scores than those the moderator listed. For example, one participant said that “I had a professor once that brought our class pizza and then took all of us that were 21 to the bars the last night of the class...we all rated him real high, even though he may not have been the greatest teacher...he knew how to make class fun.”

In the next section of the focus group discussions the moderator asked the participants questions about student characteristics that may have an effect on the

evaluation scores. The moderator asked if the participants felt they would rate a professor they felt to be more like themselves higher or lower than other professors. The participants said no, almost unanimously. One participant said “that’s almost dumb...a lot of my favorite professors are nothing like me.” Another stated that “if every professor I had was like me I think I’d do nuts.”

Next the moderator asked the participants if they would rate a professor higher or lower if they were expecting a higher grade in the class. All of the participants agreed that they would rate a professor higher if they were expecting a higher grade in the class. One participant said that “if I’m expecting a high grade in the class that means one of two things, either the professor is good and I’m learning what I’m supposed to, or the professor is easy and I’ve coasted through the course without doing much, both of those are good...of course I’m going to rate them higher.” Another participant said that they would rate a teacher higher if they were getting a higher grade because “...obviously if I’m getting a good grade they are doing what they are supposed to be doing...”

The moderator asked the participants whether or not they felt the reason they were taking the class had an affect on the rating they would give the professor. There were differing responses. Some participants felt that even if the class was required, but not in their major, and they did not like the class that would not cause them to rate the instructor lower. One participant said “its not their fault that you had to take the class and you hate it...they still are probably a very good teacher.” Another said that “if I don’t

like the class, because I had to take it to get all my credits, then I'm probably not going to like the professor either." All of the participants seemed to fit into one group or the other.

The moderator gave the participants a list of student characteristics (Appendix IV) and asked them if they felt any of them would affect the way they rated an instructor. They said none on the list would affect their rating. All of the participants agreed that those characteristics on the list had nothing to do with the teaching ability of the professor. They also agreed that none of the characteristics on the list had anything to do with even the like ability of the professor.

The final section of the focus group discussions consisted of miscellaneous questions about the student evaluations. First, the moderator asked the participants if they rate professors that entertain rather than teach higher than other professors. Most of the participants said that it depends on the teacher and the class.

- One of the participants said that "I think it depends on the teacher...some of them you like and some of them you don't... it has nothing to do with whether they entertain or teach..."
- Another participant said "it just depends on the professor, some can talk and talk and talk and it drives you crazy, others can do the same thing and you think its funny and it makes the class time go faster.. then you rate them according to that."

- Still another participant said “it doesn’t really have anything to do with entertaining or teaching...you rate them higher or lower if you like them or not.”

Next, the moderator asked the participants if they tend to focus their evaluation on teacher characteristics and what they learned in the class or the organization and structure of the class. Most of the participants said they use a mixture of both more than one or the other.

- One participant said “its not really one or the other...it takes teacher characteristics, what you learned, and the organization and structure in the class to make a good professor.”
- Another participant stated that “it takes all of that and more to be a professor.” Most of the participants agreed with both of those comments.
- One participant disagreed. They said “it all depends on structure and organization, without that you can’t learn anything anyway.”

The final question that the moderator asked the participants in the focus group discussions was “does the method used to collect the information influence the students’ evaluations of the overall teaching competence of an instructor or the quality of the course? Do you see any other alternatives to this method of collection?” The participants said that they felt that the fact that the professor being evaluated is not in the room during the evaluation process makes it easier to be honest and open in the evaluation scores. One participant said “if the professor was in the room I would probably be afraid to say what I really wanted to because he might look at it after I turned it in.” However,

most of the participants felt that the manner in which the evaluations are collected in the College of Business at Ball State University is the best way to collect the information needed in the evaluations. They felt that any other method of collection would hinder the collection of valid and reliable data.

### **Appendix III: Potential Biases**

- course difficulty
- grading leniency
- instructor popularity
- student interest in subject
- course workload
- class size/ enrollment
- required course vs. elective course

#### **Appendix IV: Student Characteristics**

- age
- sex
- college year
- academic ability
- GPA
- expected grade in the class
- reason for taking the class
- personality differences



Semester \_\_\_\_\_ Year \_\_\_\_\_ Dept \_\_\_\_\_ 5

1

**Please complete reverse side**

# STRENGTHS, WEAKNESSES, AND RECOMMENDATIONS

Please write your views of this instructor's strengths and weaknesses and recommendations for improvement. Please be specific so that your instructor can obtain meaningful feedback from your suggestions.

	1	2	3	4	5
<b>This instructor's strengths are:</b>	1 1 1 1 1 1 1 1 1	2 2 2 2 2 2 2 2 2	3 3 3 3 3 3 3 3 3	4 4 4 4 4 4 4 4 4	5 5 5 5 5 5 5 5 5
<b>This instructor's weaknesses are:</b>	1 1 1 1 1 1 1 1 1	2 2 2 2 2 2 2 2 2	3 3 3 3 3 3 3 3 3	4 4 4 4 4 4 4 4 4	5 5 5 5 5 5 5 5 5
<b>My recommendations to this instructor for teaching improvement are:</b>	1 1 1 1 1 1 1 1 1	2 2 2 2 2 2 2 2 2	3 3 3 3 3 3 3 3 3	4 4 4 4 4 4 4 4 4	5 5 5 5 5 5 5 5 5
	1 1 1 1 1 1 1 1 1	2 2 2 2 2 2 2 2 2	3 3 3 3 3 3 3 3 3	4 4 4 4 4 4 4 4 4	5 5 5 5 5 5 5 5 5

## **Appendix VI: Focus Group Participants**

Monday, March 29, 1999 at 9:30pm

Chad Hannah	284-7043
Francesca DelGuidice	287-8842
Doug Clark	288-3092
Craig Mince	621-1106
Mike Fine	284-7045
Brooke VanPutton	286-7310
Amanda Meadows	287-8145
Josh Dillon	289-6605

Wednesday, March 31, 1999 at 9:30pm

Matt Cook	741-3067 ext. 101
Bekah Hunt	287-8845
Chad Olinger	287-8751
Jamie Balhon	286-7310
Jeff Hersch	284-7045
Erin Roberts	284-3991
Scott Hagan	284-7045
Karen Gable	288-4106